

EPISEN

Ing2 2021-2022

Module BDM

# BASES DE DONNÉES MULTIDIMENSIONNELLES ET NOMADES

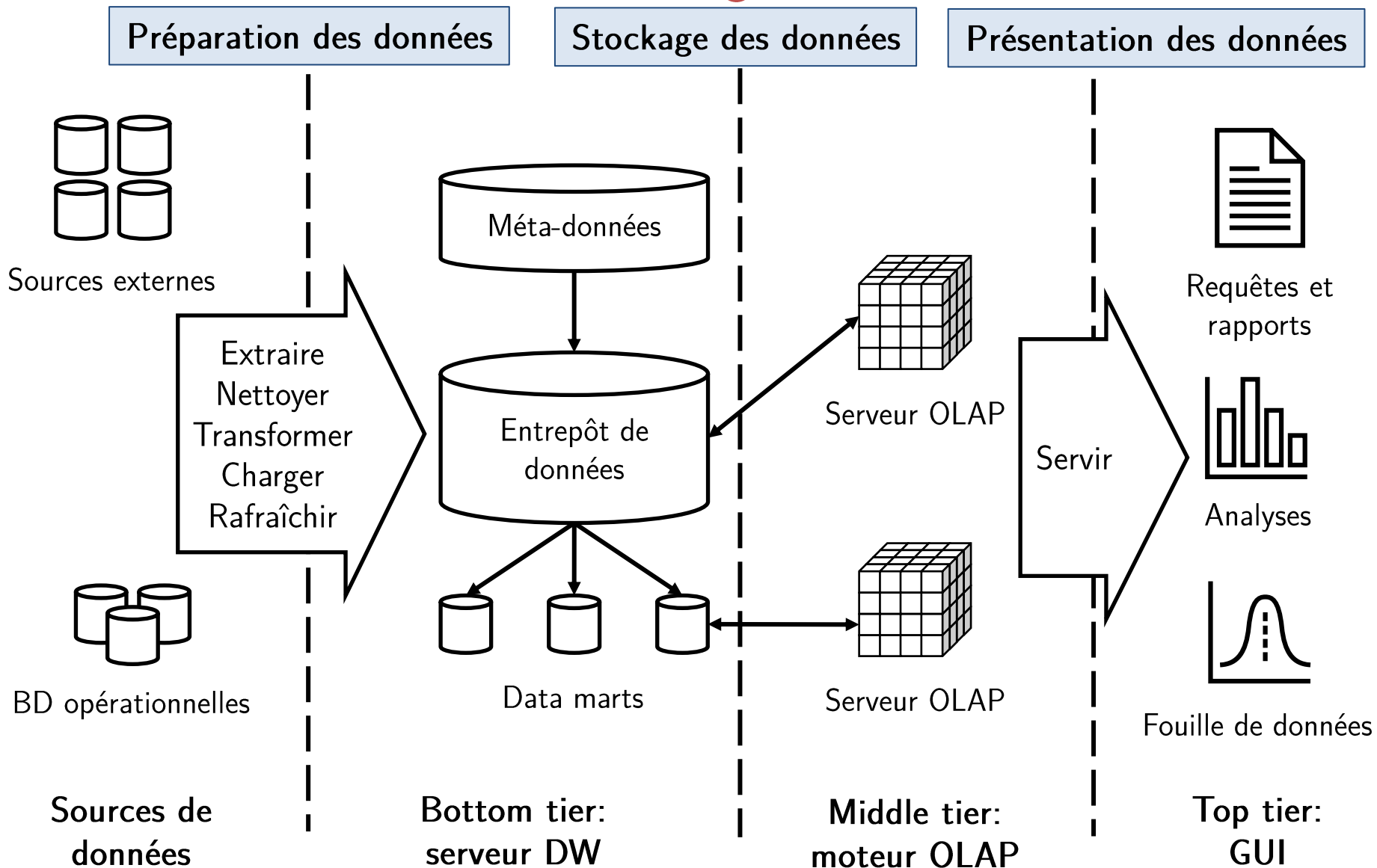
- I. Contexte
- II. Modélisation des entrepôts de données (DW)
- III. **Alimentation des entrepôts de données**
  - 1. Extraction, transformation, chargement
  - 2. Un exemple de processus E.T.L
- IV. Accès aux données de l'entrepôt
- V. Perspectives et évolution

# ALIMENTATION DES ENTREPÔTS DE DONNÉES

- I. Contexte
- II. Modélisation des entrepôts de données (DW)
- III. Alimentation des entrepôts de données
- IV. Accès aux données de l'entrepôt
  - 1. Le cube de données
  - 2. Type de cubes
  - 3. Analyse multidimensionnelle
  - 4. Outil d'analyse
- V. Perspectives et évolution

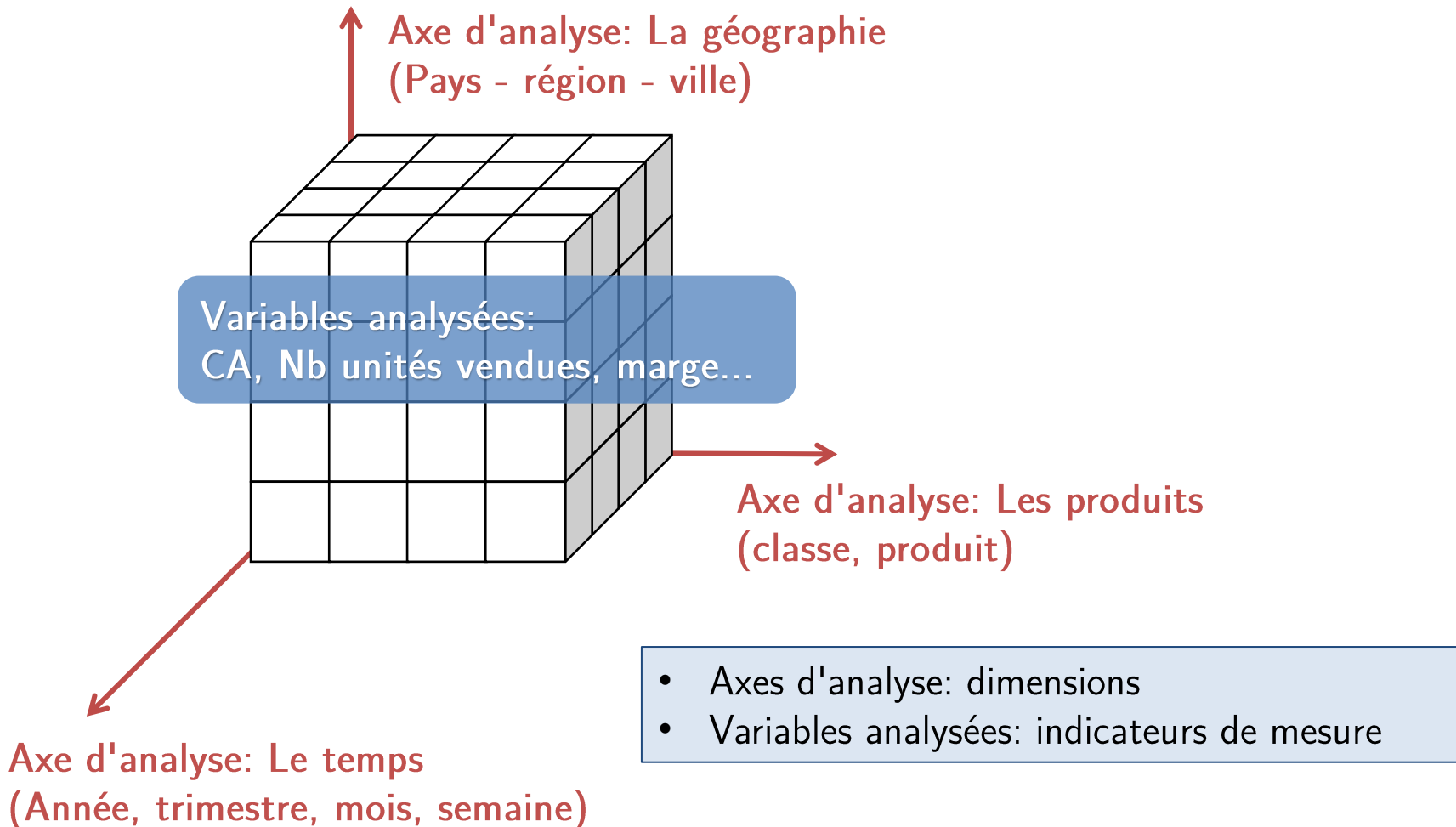
# ACCÈS AUX DONNÉES DE L'ENTREPÔT

# Architecture générale



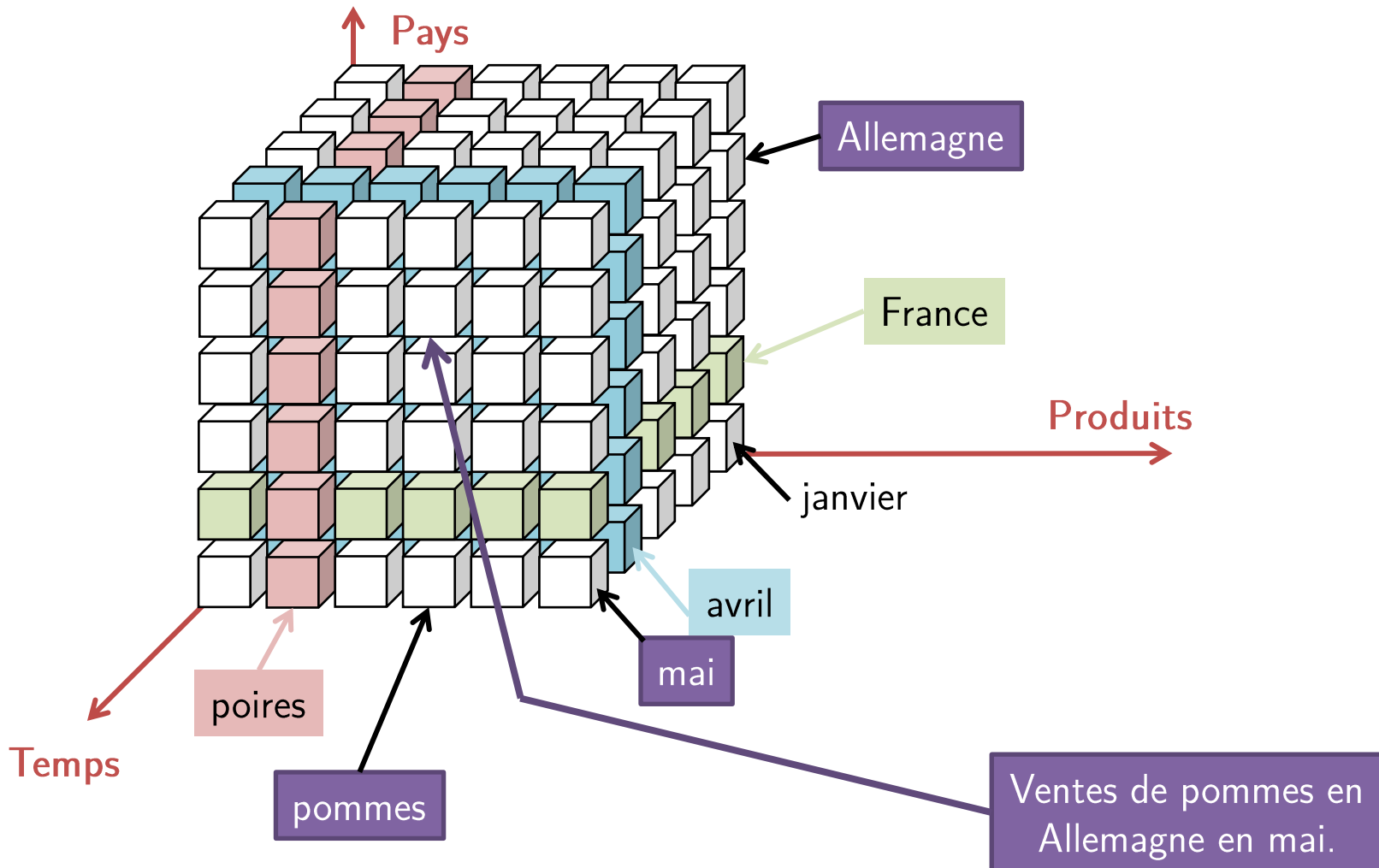
# Le cube

Un cube est défini par des dimensions ; au croisement de ces dimensions se trouvent des mesures.



# Le cube

Un cube est un ensemble de données construit à partir d'un entrepôt de données ou d'un datamart.



# OLAP & le Data Mining

- Les applications OLAP : utilisées pour analyser les performances de l'entreprise (technologie OLAP)
  - Ventes réalisées en quantité et valeur par point de vente pour chaque collection d'ouvrages?
- Le Data Mining (fouille de données) utilise des algorithmes de reconnaissance de modèles afin de détecter des **comportements particuliers**, des **corrélations** ou des **tendances** dans les données
  - Les informations obtenues servent à des fins de prédiction telles que des prévisions de ventes, segmentation de population d'individus au comportement similaire etc

# Les types de cubes

## 1. Relational OLAP (ROLAP)

- Données stockées dans une base de données relationnelle
- Un moteur OLAP permet de simuler le comportement d'un SGBD multidimensionnel
  
- Plus facile et moins cher à mettre en place
- Moins performant lors des phases de calcul
  
- Exemple de moteurs ROLAP:
  - Mondrian



# Les types de cubes

## 2. Multidimensional OLAP (MOLAP)

- un système multidimensionnel « pur » qui gère les structures multidimensionnelles natives (les cubes)
- Accès direct aux données dans le cube
  
- Plus difficile à mettre en place
- Formats souvent propriétaires ; conçus exclusivement pour l'analyse multidimensionnelle
  
- Exemple de moteurs MOLAP:
  - Microsoft Analysis Services
  - Hyperion

# Les types de cubes

## 3. Hybride OLAP (HOLAP)

- tables de faits et tables de dimensions stockées dans des SGBD relationnel (données de base)
- données agrégées stockées dans des cubes
  
- Solution hybride entre MOLAP et ROLAP
- Bon compromis au niveau coût et performance

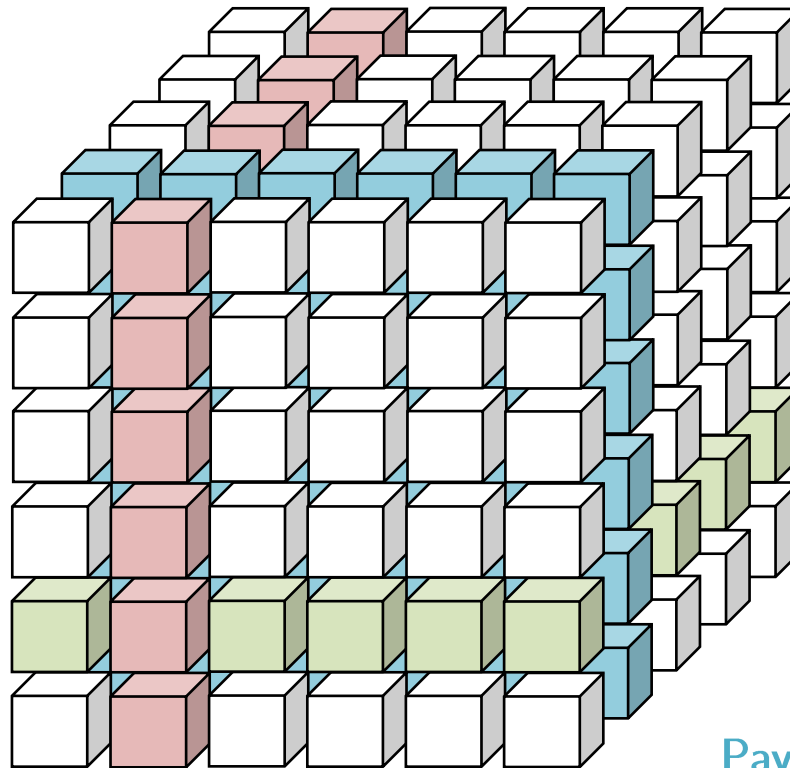
# Outils d'analyse OLAP

- Ces outils permettent de manipuler les données en effectuant des rotations de dimensions ou de hiérarchies, des forages vers le bas ou vers le haut (extensions du langage SQL).
- La manipulation peut se faire de manière graphique à l'aide de tableaux à 2 ou  $n$  dimensions
- Ces types de manipulation reposent sur de l'algèbre dite multidimensionnelle (algèbre de cubes) par opposition à l'algèbre relationnelle.

# Slice

## Définition

Une tranche (slice) est un sous-ensemble du tableau contenant toutes les cases ayant une valeur donnée pour une ou plusieurs dimensions.



Date = 1/1/22

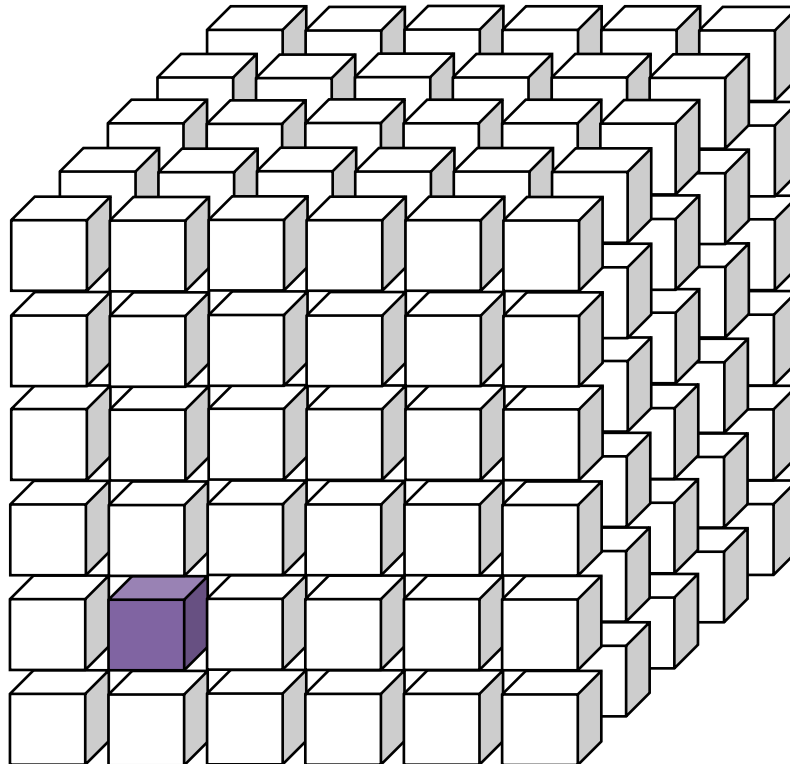
Pays = France

Produit = pomme

# Dice

## Définition

Un dé (dice) est une tranche prise sur deux dimensions ou plus.



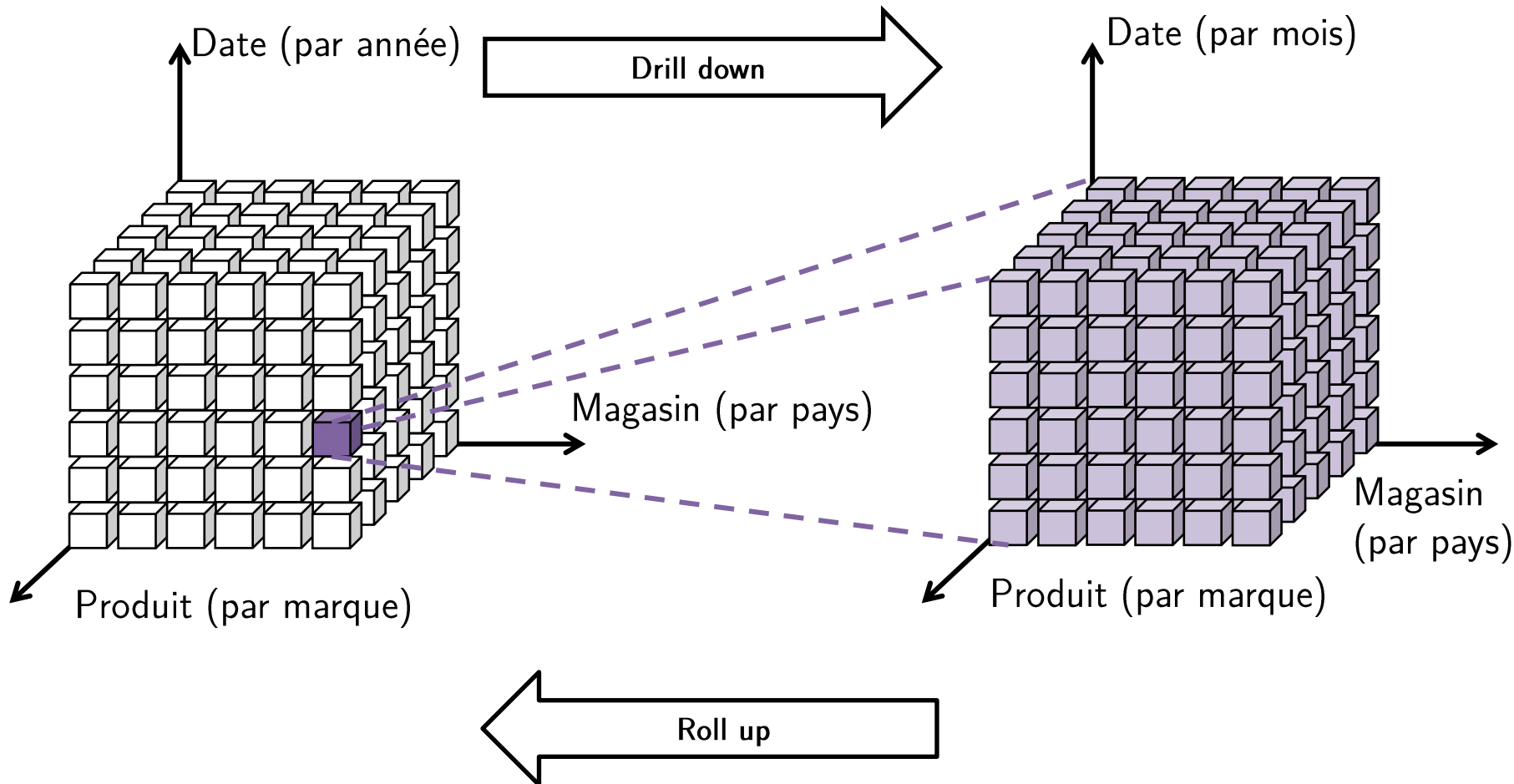
Date = 1/1/22

Produit = pomme

# Drill down/Roll up

## Définition

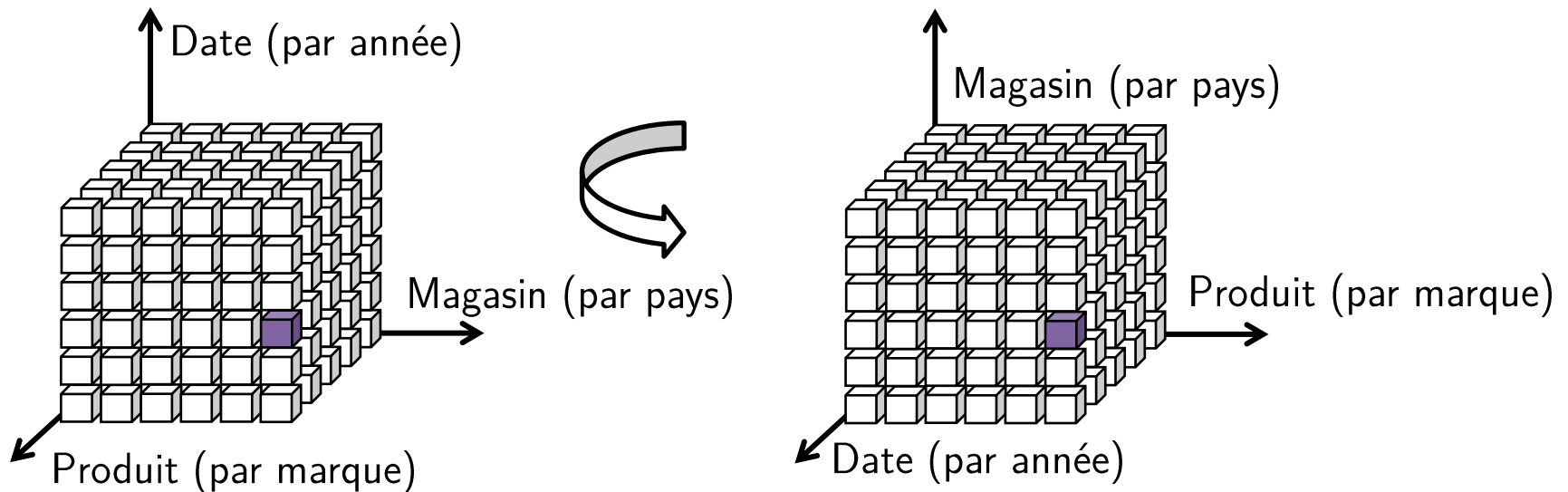
Creuser vers le haut (respectivement le bas) consiste à remplacer une dimension une autre, plus générale (respectivement plus détaillée), suivant la hiérarchie de la dimension.



# Pivot/rotate

## Définition

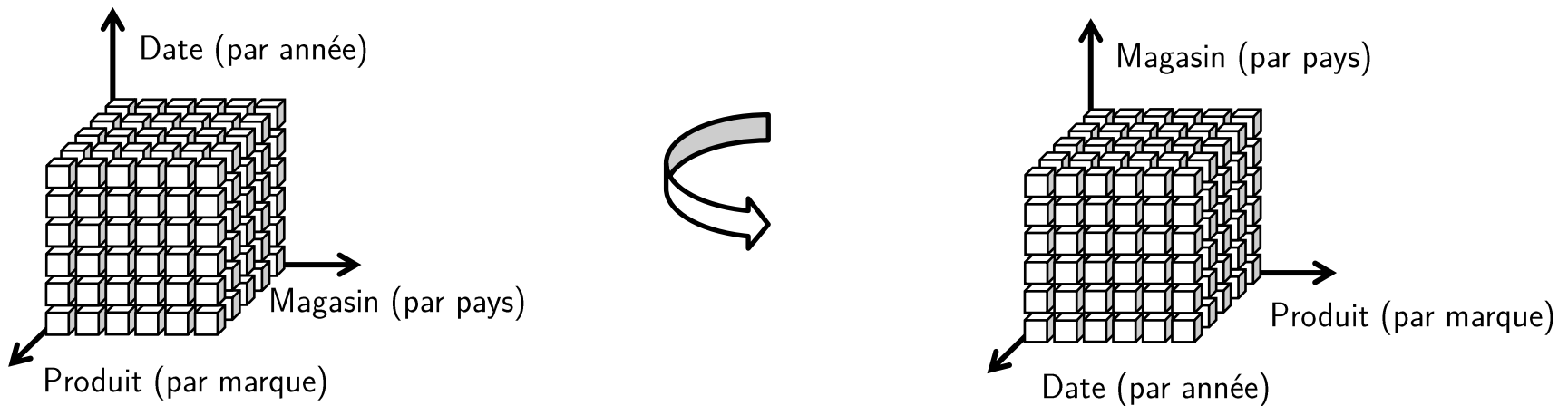
Faire pivoter/tourner un cube consiste à modifier l'ordre des dimensions, c'est-à-dire la manière dont les dimensions sont associées aux axes. Cela ne change pas les données elles-mêmes, mais plutôt l'affichage qui en est fait par l'outil.



# Pivot/rotate

## Définition

Faire pivoter/tourner un cube consiste à modifier l'ordre des dimensions, c'est-à-dire la manière dont les dimensions sont associées aux axes. Cela ne change pas les données elles-mêmes, mais plutôt l'affichage qui en est fait par l'outil.



	2020		2021	
	France	UK	France	UK
Nike	1	4	7	10
Adidas	2	5	8	11
Puma	3	6	9	12

	France			UK		
	Nike	Adidas	Puma	Nike	Adidas	Puma
2020	1	2	3	4	5	6
2021	7	8	9	10	11	12



